The noise in the circular law and the Gaussian free field

Brian Rider and Bálint Virág

February 2, 2008

Abstract

Fill an $n \times n$ matrix with independent complex Gaussians of variance $1/n$. As $n \to \infty$, the eigenvalues $\{z_k\}$ converge to a sum of an H^1-noise on the unit disk and an independent $H^{1/2}$-noise on the unit circle. More precisely, for C^1 functions of suitable growth, the distribution of $\sum_{k=1}^{n}(f(z_k) - \mathbf{E}f(z_k))$ converges to that of a mean-zero Gaussian with variance given by the sum of the squares of the disk H^1 and the circle $H^{1/2}$ norms of f. As a consequence, with p_n the characteristic polynomial, it is found that $\log|p_n| - \mathbf{E}\log|p_n|$ tends to the planar Gaussian free field conditioned to be harmonic outside the unit disk. Further, for polynomial test functions f, we prove that the limiting covariance structure is universal for a class of models including Haar distributed unitary matrices.

1 Introduction

Consider the *Ginibre ensemble*, that is the $n \times n$ random matrix in which all entries are independent complex Gaussians of mean zero and variance $1/n$. The eigenvalues $z_1, z_2, \ldots, z_n$ form a point process, a realization of which is depicted on the left of Figure 1. A first glance

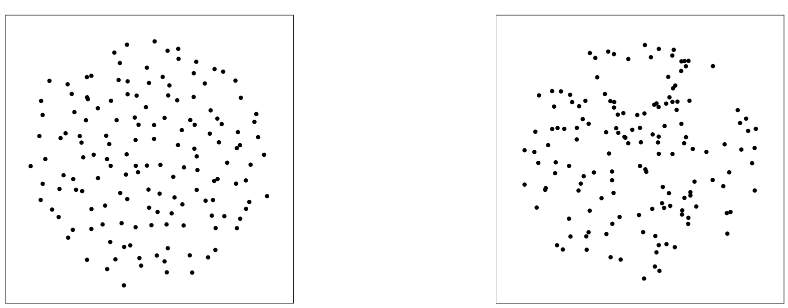

Figure 1: Ginibre eigenvalues and uniform points in $\mathbb{U}$, $n = 150$.

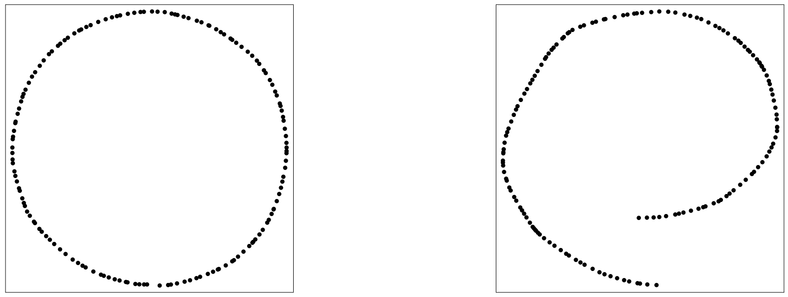

Figure 2: The running eigenvalue sums $\sum_{k=1}^{m} z_k$, $m = 1, \ldots, 150$, in which the z_k are ordered by their arguments, and the same for the independent points.

at this picture suggests that the eigenvalues are uniformly distributed in the unit disk $\mathbb{U}$. Indeed, Bai [2] has shown that, with probability one, the empirical measure converges to the uniform distribution on $\mathbb{U}$. This is the circular law of the title.

Compared to n points dropped independently and uniformly on $\mathbb{U}$ (see the right of Figure 1), the Ginibre points are clearly more regular. For example, their sum, the matrix trace, is complex Gaussian with variance 1, while for the independent points the variance is $\frac{2}{3}n$ (see Figure 2 for comparison). Our main theorem addresses this phenomenon in greater generality, rigorously verifying a prediction of Forrester [7].

Theorem 1 (Noise limit for eigenvalues). *Let $f : \mathbb{C} \to \mathbb{R}$ possess continuous partial derivatives in a neighborhood of $\mathbb{U}$, and grow at most exponentially at infinity. Then, as $n \to \infty$, the distribution of the random variable $\sum_{k=1}^{n} (f(z_k) - \mathbf{E}f(z_k))$ converges to a normal with variance $\sigma_f^2 + \tilde{\sigma}_f^2$, where*

$$\sigma_f^2 = \frac{1}{4\pi}\|f\|_{H^1(\mathbb{U})}^2 = \frac{1}{4\pi}\int_{\mathbb{U}} |\nabla f|^2 d^2 z$$

is the squared Dirichlet (H^1) norm of f on the unit disk, and

$$\tilde{\sigma}_f^2 = \frac{1}{2}\|f\|_{H^{1/2}(\partial\mathbb{U})}^2 = \frac{1}{2}\sum_{k\in\mathbb{Z}} |k||\hat{f}(k)|^2$$

is the $H^{1/2}$-norm on the unit circle $\partial\mathbb{U}$, where $\hat{f}(k) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-ik\theta}d\theta$ is the k-th Fourier coefficient of f restricted to $|z| = 1$.

We mention that the centralizer $\sum_{k=1}^{n} \mathbf{E}f(z_k)$ is easily replaced by $n \times \frac{1}{\pi}\int_{\mathbb{U}} f(z)d^2z$ with no change to the outcome. More interesting is the covariance structure that appears above, being composed of two independent terms: one for the boundary and one for the bulk. An

interpretation of this limiting noise is provided upon considering the characteristic polynomial $p_n(z) = \prod_{k=1}^n (z - z_k)$ which has fluctuations described in terms of the Gaussian free field.

The planar Gaussian free field (GFF) is a model which has received considerable recent attention as the scaling limit of uniformly random (discrete) $\mathbb{R}^2 \to \mathbb{R}$ surfaces, though it apparently has not previously been connected with any matrix model. While the realizations of the GFF fluctuate too wildly to allow it to be defined pointwise, it may be defined as a random distribution h as follows: for all suitable test functions f, the random variable $\langle f, h \rangle_{H^1(\mathbb{C})}$ is Gaussian with variance $\|f\|^2_{H^1(\mathbb{C})} = \int_{\mathbb{C}} |\nabla f|^2 \, d^2 z$. In more classical language this is the Gaussian Hilbert space for $H^1(\mathbb{C})$, identifying the GFF as the most natural 2-dimensional analogue of Brownian motion.

As the Ginibre eigenvalues accumulate in the unit disk, $\log |p_n(z)|$ tends to be harmonic away from the unit disk as $n \to \infty$. Thus the following version of the GFF appears naturally in the limit.

Corollary 2 (The Gaussian free field limit). *For the Ginibre characteristic polynomial* $p_n(z) = \prod_{k=1}^n (z - z_k)$*, let*

$$h_n(z) = \log |p_n(z)| - \mathbf{E} \log |p_n(z)|.$$

Then h_n converges weakly without normalization to h^*, the planar Gaussian free field conditioned to be harmonic outside the disk. More precisely, for functions f as above we have

$$\int h_n(z) f(z) \, d^2 z \Rightarrow \int h^*(z) f(z) \, d^2 z$$

in distribution, and the same holds for the joint distribution for the integrals against multiple test functions $f_1, \dots, f_k$.

Next, by the Cramér-Wold device, a version of Theorem 1 holds for complex valued test functions f of like regularity and growth. In this case the H^1 noise term is replaced by the L^2 norm of the $\bar{\partial} = (1/2)(\partial_x + \partial_y)$ derivative of f. Then, if it is further assumed that f is complex analytic in a neighborhood of $\overline{\mathbb{U}}$, the covariance structure simplifies significantly:

$$\sum_{k=1}^n (f(z_k) - \mathbf{E} f(z_k)) \Rightarrow \mathcal{N}_{\mathbb{C}}\Big(0, \frac{1}{\pi} \|f'\|^2_{L^2(\mathbb{U})}\Big) = \mathcal{N}_{\mathbb{C}}\Big(0, \|f\|_{H^{1/2}(\partial U)}\Big),$$

where $\mathcal{N}_{\mathbb{C}}$ denotes complex Gaussian distribution with given mean and variance. This should be compared with the result of Rider and Silverstein [18], where the same central limit

theorem is proved for random matrices with more general than $\mathcal{N}_{\mathbb{C}}(0,1)$ entries, but where the test functions need to be analytic throughout $|z| \le 4$.

Perhaps surprisingly, exactly the same limiting covariance structure arises from eigenvalues of Haar distributed unitary matrices (Diaconis and Evans, [6]). While the difference of the two limiting noises seems very clear (one lives on the disk and one on the circle), analytic functions do not see this. We show that this is not a coincidence by proving a universality result.

The important point is that both the Ginibre eigenvalues and those of Haar $U(n)$ comprise *symmetric polynomial projection (SPP) processes* in $\mathbb{C}$. An SPP process is a determinantal process defined by a pair (μ, n), where μ is a rotationally invariant probability measure on $\mathbb{C}$ and n is the number of points, as follows. For $k = n$ the joint intensity $p(z_1, \ldots, z_k)$ with respect to μ^k is given by $\det(K(z_i, z_j)_{1\le i,j\le k})$, for K the integral projection kernel to the subspace of polynomials of degree less than n in $L^2(\mathbb{C}, \mu)$. In fact, as soon as this holds for $k = n$, it holds for any k, see for example [11].

Theorem 3 (Universal noise limit for analytic polynomial test functions). *Consider a sequence of SPP processes defined by (μ_n, n). Assume that for all integers m we have*

$$M(n, 2n + 2m)/M(n, 2n) \to 1, \qquad \text{where } M(n,k) = \int_{\mathbb{C}} |z|^k d\mu_n(z),$$

as $n \to \infty$. Then, for f any polynomial in z,

$$\sum_{k=1}^{n} \left(f(z_k) - \mathbf{E} f(z_k)\right) \Rightarrow \mathcal{N}_{\mathbb{C}}\Big(0, \frac{1}{\pi}\|f'\|^2_{L^2(\mathbb{U})}\Big),$$

in distribution.

The proof of Theorem 3 relies on the close connection between SPP processes and Schur functions. For unitary matrices, this connection is a consequence of Frobenius duality in representation theory (see [6]); in the last part of the paper we show that they are useful in a more general setting. For the Ginibre ensemble $\mu_n = \mathcal{N}_{\mathbb{C}}(0, 1/n)$, while for unitary eigenvalues μ_n is uniform measure on the unit circle. When μ_n is uniform measure in the unit disk we get the (truncated) Bergman ensemble, for which the $n \to \infty$ limit is a *discrete* point process in $\mathbb{U}$ that agrees with the zeros of the random power series with i.i.d. complex Gaussian coefficients, see [11]. Each of these ensembles satisfies the moment condition laid out in the statement, and in each case the result actually holds for f analytic.

Methods. Central limit theorems of the above type have been studied in the random matrix theory community for matrix ensembles Hermitian ensembles, e.g., [13, 9, 3, 1], and

ensembles connected with the classical compact groups, e.g., [12, 23, 6, 25]. In all of these cases, the eigenvalues form a one-dimensional point process either on the real line or on the unit circle.

The situation for two-dimensional eigenvalue processes is more complicated; indeed, it seems that this paper is the first to prove a central limit theorem in two dimensions for general (not necessarily analytic) test functions. For the one-dimensional case, variants Wigner's method of moments argument has been refined and extended to give central limit theorems in several settings. However, this argument does not work for non-analytic test functions, as such functions are not expressible as a trace of a matrix polynomial. While analyticity in one dimension is essentially a smoothness assumption, in two-dimensions it is much more restrictive. This makes the proof of the general circular law [2] more difficult than that of the semicircle law, and explains why the corresponding general central limit theorem is still open (again, see [18] for analytic test functions).

Our proof is based on the determinantal structure of the point process and a cumulant formula for determinantal processes introduced by Costin-Lebowitz [5] and generalized by Soshnikov [22, 23, 24]. In the present setting, it allows us to establish a connection between joint cumulants of monomial test functions in z and $\bar{z}$ and a class of combinatorial objects which we call rotary flows.

When the associated matrix integrals can be evaluated by separate methods, then such a connection can be used to count intractable combinatorial objects or evaluate difficult combinatorial sums arising, for example, in statistical physics. Here the cumulants become combinatorial sums over rotary flows. Fortunately, we are able to establish (Section 5) conditions under which these combinatorial sums vanish as $n \to \infty$, without explicitly computing them for any finite n case. As it turn out, this is sufficient to conclude the asymptotic joint normality of monomial statistics. This result is then lifted from polynomial to general C^1 functionals by a concentration estimate, yielding Theorem 1. The sum of these steps occupies Sections 4 through 7. Theorem 3 is proved in Section 8, while, after some preliminaries in the next section, the connection to the GFF is detailed in Section 3.

The arguments introduced here work in the general rotational invariant setting of SPP processes. However, beyond analytic test functions, the structure of the limit depends intimately on the process in question. In this paper we concentrate on the Ginibre ensemble.

2 Facts about the Ginibre ensemble

The non-Hermitian matrix ensemble of independent complex Gaussian entries is named for Ginibre who discovered [8] that the joint density of eigenvalues, z_1 through z_n, is given by

$$d\mu_n(z_1)\cdots d\mu_n(z_n)\,\frac{1}{\mathcal{Z}_n}\prod_{j<k}|z_j-z_k|^2. \tag{1}$$

Here $d\mu_n = \frac{n}{\pi}e^{-n|z|^2}d^2z$ is complex Gaussian measure of variance $1/n$, and $\mathcal{Z}_n < \infty$ is a normalizing constant.

A description equivalent to (1) is to say that the eigenvalues of the Ginibre ensemble make up the determinantal process tied to the projection onto the subspace of polynomials of degree less than n in $L^2(\mathbb{C},\mu_n)$; the projection kernel with respect to μ_n being

$$K_n(z,\bar{w}) = \sum_{\ell=0}^{n-1}\frac{n^\ell}{\ell!}z^\ell\bar{w}^\ell.$$

That is, all finite dimensional correlation functions, or joint intensities, are expressed in terms of determinants of the kernel K_n: for $B_1,\ldots,B_k$ any mutually disjoint family of Borel subsets of $\mathbb{C}$,

$$\mathbf{E}\Big[\prod_{\ell=1}^{k}\#\{k : z_k\in B_\ell\}\Big] = \int_{\prod_{\ell=1}^k B_\ell}\det\Big(K_n(z_i,z_j)\Big)_{1\le i,j\le k}d\mu_n(z_1)\cdots d\mu_n(z_k). \tag{2}$$

One then refers to the process (K_n,μ_n). Further background information on determinantal processes may be found in [11] and [22]. Immediate from either (1) or (2) is that the eigenvalues repel each other; Figure 1 provides a somewhat dramatic snapshot of this fact.

As the dimension tends to infinity, the reader is invited to check by a simple calculation that the mean density of Ginibre eigenvalues,

$$\mathbf{E}\Big[\frac{1}{n}\sum_{k=1}^{n}\delta_{z_k}(z)\Big] = \frac{1}{n}K_n(z,\bar{z})d\mu_n(z),$$

tends (weakly) to the uniform measure on the the unit disk $\mathbb{U}$. This is a particularly simple instance of the circular law about which we describe fluctuations.

3 The Gaussian free field

For a domain $\mathbb{D}\subset\mathbb{C}$, let $H^1 = H^1(\mathbb{D})$ denote the Hilbert space completion of smooth functions with compact support in $\mathbb{D}$ with respect to the gradient inner product $\langle f,g\rangle_{H^1(\mathbb{D})} =$

$\langle f,g\rangle_{H^1(\mathbb{D})} = \int_{\mathbb{D}} \nabla f \cdot \nabla g \, d^2 z$. Note that the H^1-norm is invariant under conformal transformations and complex conjugation.

The Gaussian free field (GFF) on $\mathbb{D}$ is defined as the Gaussian Hilbert space connected with this norm. It may also be interpreted as the random distribution h such that for all $f \in H^1(\mathbb{D})$ the random variable $\langle f, h\rangle_{H^1(\mathbb{D})}$ is centered normal with

$$\operatorname{Var}\left(\langle f, h\rangle_{H^1(\mathbb{D})}\right) = \|f\|^2_{H^1(\mathbb{D})}.$$

For more detailed information concerning the GFF, the paper [19] is recommended.

In contrast, the H^1-noise on $\mathbb{D}$ may be defined as the Gaussian random distribution $\tilde{h}$ for which $\langle f, \tilde{h}\rangle_{L^2}$ is centered normal and

$$\operatorname{Var}\left(\langle f, \tilde{h}\rangle_{L^2}\right) = \|f\|^2_{H^1(\mathbb{D})}.$$

The two are related by noting that

$$\langle f, h\rangle_{H^1} = \langle -\triangle f, h\rangle_{L^2} = \langle f, -\triangle h\rangle_{L^2}.$$

That is, formally there is the identity in law $\tilde{h} = -\triangle h$.

Planar GFF harmonic outside $\mathbb{U}$ and Corollary 2

The space $H^1(\mathbb{C})$ decomposes into three orthogonal parts (see [19]): $H^1(\mathbb{C}) = A_I \oplus A_O \oplus A_H$. A_I is the closure (in H^1) of smooth functions supported in a compact set of the open unit disk. A_O is the similar closure of functions supported in a compact set in the complement of the closed unit disk. Last, A_H is the subspace spanned by continuous functions which are harmonic on the complement of the unit circle.

We can now define the GFF h on $\mathbb{C}$ conditioned to be harmonic exterior to $\mathbb{U}$ as the projection $\mathcal{P}_{IH}$ of the standard planar GFF to $A_I \oplus A_H$ in the distributional sense: for $f \in H^1$, set $\langle f, \mathcal{P}_{IH} h\rangle = \langle \mathcal{P}_{IH} f, h\rangle$. Note that for $f \in H^1$ and continuous, $\mathcal{P}_H f$ is the harmonic extension of the values of f on the unit circle to the whole plane. That is, the resulting function is harmonic at infinity, or $(\mathcal{P}_H f)(1/z)$ is harmonic near the origin.

The proof of Corollary 2 requires one preliminary observation.

Lemma 4. *For $\mathcal{P}_{IH}$ the projection unto the $A_I \oplus A_H$ it holds that*

$$\|\mathcal{P}_{IH} f\|^2_{H^1(\mathbb{C})} = \|f\|^2_{H^1(\mathbb{U})} + \frac{1}{2}\|\mathcal{P}_H f\|^2_{H^1(\mathbb{C})}.$$

For f continuous in the neighborhood of ∂U the second term satisfies

$$\frac{1}{2}\|\mathcal{P}_H f\|^2_{H^1(\mathbb{C})} = \|\mathcal{P}_H f\|^2_{H^1(\mathbb{U})} = \pi\|f\|^2_{H^{1/2}(\partial\mathbb{U})}. \tag{3}$$

Proof. If $f \in H^1(\mathbb{C})$, there exists a function $g \in A_O$ so that $s = f + g$ is symmetric with respect to inversion, *i.e.*, $s(z) = s(1/\bar{z})$. With $\|\cdot\|$ denoting $H^1(\mathbb{C})$ unless specified otherwise, we write

$$\|s\|^2 = \|\mathcal{P}_I s\|^2 + \|\mathcal{P}_O s\|^2 + \|\mathcal{P}_H s\|^2.$$

By symmetry and the conformal and conjugation invariance of H^1, we have that $\|\mathcal{P}_I s\| = \|\mathcal{P}_O s\|$, and also $\|s\|^2_{H^1(\mathbb{U})} = \|s\|^2_{H^1(\mathbb{C}\setminus\mathbb{U})} = \|s\|^2/2$. Hence,

$$\|\mathcal{P}_I s\|^2 = \|s\|^2_{H^1(\mathbb{U})} - \frac{1}{2}\|\mathcal{P}_H s\|^2,$$

and we conclude that

$$\|\mathcal{P}_{IH} f\|^2 = \|\mathcal{P}_{IH} s\|^2 = \|s\|^2_{H^1(\mathbb{U})} + \frac{1}{2}\|\mathcal{P}_H s\|^2 = \|f\|^2_{H^1(\mathbb{U})} + \frac{1}{2}\|\mathcal{P}_H f\|^2$$

since $g = s - f \in A_O$. This proves the first claim.

The first equality in (3) follows from symmetry and conformal (inversion) and conjugation invariance. Note that all norms in (3) depend only on the values of f on ∂U, so we may assume that f is harmonic in $\mathbb{U}$, and drop the projection. Harmonic functions in $\mathbb{U}$ are spanned by the real and imaginary parts of z^k, so it suffices to check that for two elements g_1, g_2 of this spanning set

$$\langle g_1, g_2 \rangle_{H^1(\mathbb{U})} = \pi \langle g_1, g_2 \rangle_{H^{1/2}(\partial\mathbb{U})},$$

which is a simple computation. □

Proof of Corollary 2. By Lemma 4 the limiting variance in Theorem 1 can be written as

$$\frac{1}{4\pi}||f||^2_{H^1(\mathbb{U})} + \frac{1}{2}||f||^2_{H^{1/2}(\partial\mathbb{U})} = \frac{1}{4\pi}\|\mathcal{P}_{IH} f\|^2_{H_1(\mathbb{C})}.$$

Now recall $p_n(z) = \prod_{k=1}^n (z - z_k)$, the Ginibre characteristic polynomial, and set

$$\tilde{h}_n(z) = \frac{1}{2\pi} \log |p_n(z)| - \frac{1}{2\pi}\mathbf{E} \log |p_n(z)|.$$

Since $\frac{1}{2\pi}\triangle \log |z| =$ the delta measure at the origin in distribution, $h_n(z) = \triangle \tilde{h}_n(z)$ equals the centered counting measure of eigenvalues in the same sense. Thus, by Theorem 1, we have that for any f which is once differentiable in a neighborhood of $\mathbb{U}$, the random variable $\langle f, \tilde{h}_n \rangle_{L^2(\mathbb{U})}$ is asymptotically normal with variance $\frac{1}{4\pi}\|\mathcal{P}_{IH} f\|^2_{H^1(\mathbb{C})}$ as $n \to \infty$. In this way we can identify the distributional limit of $\tilde{h}_n(z)$ as the planar Gaussian free field, conditioned to be harmonic outside the unit disk. □

4 Cumulants, polynomial statistics and rotary flows

For any real-valued random variable X the cumulants, $\mathcal{C}_k(X), k = 1, 2, \ldots,$ are defined by the expansion

$$\log \mathbf{E}[e^{itX}] = \sum_{k=1}^{\infty} \frac{(it)^k}{k!} \mathcal{C}_k(X), \tag{4}$$

and carry information in the same manner as the moments of X. Important here is the fact that the variable X is Gaussian if and only if $\mathcal{C}_k(X) = 0$ for all $k > 2$.

Now, as it seems to have been first observed by Costin and Lebowitz in [5], the cumulants of linear statistics in (any) determinantal point process have a particularly nice form. Staying in $\mathbb{C}$, choose a kernel K and measure μ so that K defines a self-adjoint integral operator $\mathcal{K}$ on $L^2(\mathbb{C}, d\mu)$. Then, if $\mathcal{K}$ is locally trace class with all eigenvalues in $[0, 1]$, (K, μ) determines a determinantal point process (see [11], [15] or [22]). That is, there is a process with all k-point intensities satisfying the identity (2). With $X(g) = \sum g(z_k)$ and $\{z_k\}$ the points of (K, μ), the k-th cumulant is written,

$$\begin{aligned} \mathcal{C}_k(X(g)) \quad &= \quad \sum_{m=1}^{k} \frac{(-1)^{m-1}}{m} \sum_{\substack{k_1+\cdots+k_m=k \\ k_1\geq 1,\ldots k_m\geq 1}} \frac{k!}{k_1! \cdots k_m!} \\ &\quad \times \int_{\mathbb{C}^k} \Bigl(\prod_{\ell=1}^{m} (g(z_\ell))^{k_\ell}\Bigr) K(z_1, \bar{z}_2) K(z_2, \bar{z}_3) \cdots K(z_m, \bar{z}_1) d\mu(z_1) \cdots d\mu(z_m). \end{aligned} \tag{5}$$

Along with [5], the above has been put to important use in [23], [24].

For what we do here it will be convenient to cast the cumulants somewhat differently. First let $[k] = \{1, \ldots, k\}$. For a function $\sigma : [k] \to [m]$, and $f \in \mathcal{G}^k$, where $\mathcal{G}$ is an algebra real-valued functions, define $\sigma f \in \mathcal{G}^m$ by

$$(\sigma f)_j(z) = \prod_{i:\, \sigma(i)=j} f_i(z).$$

Fix a functional $\Phi_m : \mathcal{G}^m \to \mathbb{R}$ for each $1 \leq m \leq k$. For $f \in \mathcal{G}^k$ define

$$\Upsilon_{k,m}(\Phi_m, f) = \sum_{\sigma[m] \twoheadrightarrow [k]} \Phi_m(\sigma f),$$

where the sum is over all surjections (denoted by $\twoheadrightarrow$). Further define

$$\Upsilon_k(\Phi, f) = \sum_{m=1}^{k} \frac{(-1)^{m-1}}{m} \Upsilon_{k,m}(\Phi_m, f). \tag{6}$$

Then, for the determinantal process (K, μ), we set

$$\Phi_m(f_1, \dots, f_m) = \int f_1(z_1) \cdots f_m(z_m) K(z_1, \bar{z}_2) K(z_2, \bar{z}_3) \cdots K(z_m, \bar{z}_1)\, d\mu(z_1) \cdots d\mu(z_m),$$

and replace the formula (5) for the k-th cumulant with the equivalent,

$$\mathcal{C}_k(X(g)) = \Upsilon_k(\Phi, (g, g, \dots, g)). \tag{7}$$

This new expression, as a sum over surjections rather than over ordered partitions, lends itself to the combinatorial approach we take. In the following, it is explained how the cumulant formulas, (6) through (7), simplify in the case that the determinatal process in question has a rotation invariant reference measure μ in $\mathbb{C}$ and the linear statistic considered is a (weighted) polynomial in $z, \bar{z}$. The next section (Section 5) establishes when the limiting combinatorial sums vanish. Finally, these facts are put together to compute the asymptotics of the cumulants of polynomial statistics in the Ginibre ensemble, establishing the CLT in that case.

Rotation invariance and rotary flows

The following identities hinge on the rotation invariance of Ginibre ensemble. To drive this point home, we set things up for general determinantal point processes which share this feature.

Begin with a radially symmetric reference measure μ in the complex plane normalized so that $\mu(\mathbb{C}) = 1$, and define the kernel,

$$K(z, \bar{w}) = \sum_{\ell=0}^{L} \lambda_\ell \, c_\ell \, (z\bar{w})^\ell, \tag{8}$$

where it is assumed that $\lambda_\ell \in [0, 1]$, $\int_{\mathbb{C}} |z|^{2\ell} d\mu = c_\ell^{-1} < \infty$, while $L = \infty$ is allowed. Again, the results of [15] or [22] (see also [11]) imply that (K, μ) define appropriate joint intensities.

Now return to the formula (7) for the k-th cumulant of the linear functional $X(g) = \sum g(z_k)$:

$$\mathcal{C}_k(X(g)) = \Upsilon_k(\Phi, (g, g, \dots, g)).$$

Since $\Upsilon(\Phi, (f_1, \dots, f_k))$ is a k-linear symmetric functional, it follows that if g is a polynomial, $g = \sum a_k f_k$, then $\mathcal{C}_k(X(g))$ is a sum of terms of the form $\Upsilon(\Phi, (f_1, \dots, f_k))$ in the monomials f_k. Our goal is to understand the conditions which will make $\Upsilon_k(\Phi, (f_1 \dots, f_k))$ vanish (or vanish in the limit of some parameter, soon to be the dimension for Ginibre).

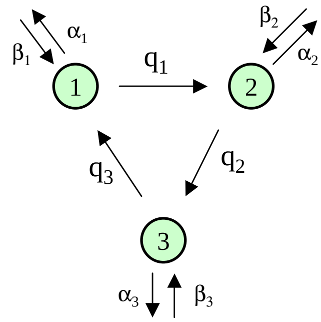

Figure 3: A rotary flow: a cumulant term that does not vanish

Looking to take advantage of the radial symmetry, we fix a simple test function $\zeta(z)$, a product of z and a radially symmetric weight function. Then taking the monomials f_j to be powers of ζ and its conjugate,

$$f_j(z) = (\zeta(z))^{\alpha_j}\overline{(\zeta(z))^{\beta_j}},$$

we expand the corresponding integral to find that

$$\Phi_m(f_1, \ldots, f_m) = \sum_{q_1,\ldots,q_m=0}^{n} \int \prod_{j=1}^{m} \zeta_j(z_j)^{\alpha_j}\overline{\zeta_j(z_j)^{\beta_j}}\lambda j c_j(z_j\bar{z}_{j+1})^{q_j}\, d\mu(z_1)\cdots d\mu(z_m).$$

Each term in the above sum can be thought of as a flow network; the vertices are $1 \ldots m$, and the directed edges are $(j, j+1)$, $j = 1 \ldots m$ (see Figure 3). The exponent of z_j in the term is $\alpha_j + q_j$, and the exponent of $\bar{z}_j$ is $\beta_j + q_{j-1}$; these have to agree or else the integral over z_j will vanish (granted by the rotation invariance of the reference measure). In flow network terminology, the node law has to hold (total inflow=total outflow), see figure 3. This implies that setting

$$\begin{aligned} \gamma_j &= \beta_j - \alpha_j, \qquad j = 1, \ldots, m, \\ \eta_j &= \gamma_1 + \ldots + \gamma_j, \end{aligned}$$

and also $\ell = q_m$, it is required that

$$\begin{aligned} q_j &= \ell + \eta_j, \qquad\qquad j = 1, \ldots, m, \\ \alpha_1 + \ldots + \alpha_m &= \beta_1 + \ldots + \beta_m, \end{aligned} \tag{9}$$

for the integral not to vanish. We refer to γ, β, η, q satisfying these conditions as a **rotary**

flow. Denoting

$$\begin{aligned} M(q) &= \int_{\mathbb{C}} |z|^q d\mu(z) = c_{q/2}^{-1}, \\ M(q,\kappa) &= \int_{\mathbb{C}} |z|^q |\zeta(z)|^\kappa d\mu(z), \end{aligned}$$

we find the following master formula,

$$\Phi_m(f_1, \ldots, f_m) = \sum_{\ell=-\eta_{min}}^{n-1-\eta_{max}} \prod_{j=1}^{m} \lambda_{\ell+\eta_j} \frac{M(2\ell + 2\eta_j - \gamma_j, \alpha_j + \beta_j)}{M(2\ell + 2\eta_j)}. \tag{10}$$

for the basic cumulant term.

In the flow language, the conditions (9) say that for each edge $(j, j+1)$ in the cycle we get a term

$$\lambda_{q_j} c_{q_j} = \lambda_{q_j} M(2q_j)^{-1} = \lambda_{\ell+\eta_j} M(2\ell + 2\eta_j)^{-1}.$$

For the vertex j the exponent of the $|z_j|$ factors coming from the kernels $K(\cdot,\cdot)$ equals the total flow within the cycle in or out of j:

$$q_{j-1} + q_j = 2\ell + \eta_{j-1} + \eta_j = 2\ell + 2\eta_j - \gamma_j.$$

The total exponent of $|\zeta(z_j)|$ is the total external flow in or out of j, which equals $\alpha_j + \beta_j$.

5 Combinatorial identities related to rotary flows

The main goal of this section is to establish general conditions under which the combinatorial sums over rotary flows arising in connection with the limits of joint cumulants vanish.

Let V be a vector space; we will use $V = \mathbb{R}$ unless otherwise specified. For a function $\sigma : [k] \to [m]$, and $\alpha \in V^k$ define $\sigma\alpha \in V^m$ by

$$(\sigma\alpha)_j = \sum_{i:\, \sigma(i)=j} \alpha_i.$$

Fix a function $\varphi_m : V^m \to \mathbb{R}$ for each $1 \le m \le k$. Define

$$\Lambda_{k,m}(\varphi_m, \alpha) = \sum_{\sigma:[m]\twoheadrightarrow[k]} \varphi_m(\sigma\alpha),$$

with the sum over all surjections, as well as

$$\Lambda_k(\varphi, \alpha) = \sum_{m=1}^{k} \frac{(-1)^{m-1}}{m} \Lambda_{k,m}(\varphi_m, \alpha).$$

The connection to the cumulant expressions above should be clear. By a slight abuse of notation, we will use the shorthand $\varphi = x^p$ for the collection of functions $\varphi_m = \alpha_1^p + ... + \alpha_m^p$. Also let

$$\Lambda_{k,m}(1) = |\{\sigma : [m] \twoheadrightarrow [k]\}| = \Lambda_{k,m}(x, (1, 0, \dots, 0)),$$

and introduce the shorthand $\Lambda_k(1) = \Lambda_k(x, (1, 0, \dots, 0))$.

Lemma 5. *For linear φ it holds that,*

$$\Lambda_k(x, \alpha) = \begin{cases} \alpha_1, & \text{if } k = 1, \\ 0, & \text{if } k \geq 2. \end{cases}$$

Proof. The case $k = 1$ is clear, so assume $k \geq 2$. Each $\varphi_k(\alpha) = \alpha_1 + \ldots + \alpha_k$, which we denote by s. We expand each $\Lambda_{k,m}$ according to whether $\sigma(k)$ has a unique pre-image. This produces

$$\Lambda_{k,m}(x, \alpha) = sm\Lambda_{k-1,m-1}(1) + sm\Lambda_{k-1,m}(1),$$

and so the expression for $\Lambda_k(x, \alpha)$ is just a telescoping sum. ∎

Lemma 6. *In the quadratic case we have*

$$\Lambda_k(x^2, \alpha) = \begin{cases} \alpha_1^2, & \text{if } k = 1, \\ 2\alpha_1\alpha_2, & \text{if } k = 2, \\ 0, & \text{if } k \geq 3. \end{cases}$$

Proof. Again assume $k \geq 2$. Let $\alpha^- = (\alpha_1, \dots, \alpha_{k-1})$, and let $s^- = \alpha_1 + \cdots + \alpha_{k-1}$. We expand $\Lambda_{k,m}$ in terms of α_k, first considering the case where $\sigma(k)$ has a unique pre-image k. Summing over values of $\sigma(k)$ we find

$$k\left(\Lambda_{k-1,m-1}(x^2, \alpha^-) + \alpha_k^2\Lambda_{k-1,m-1}(1)\right).$$

Now take the case when $\sigma(k)$ has more than one pre-image. Expanding each quadratic term yields

$$m\Lambda_{k-1,m}(x^2, \alpha^-) + m\alpha_k^2\Lambda_{k-1,m}(1) + \sum_{\sigma:[k]\twoheadrightarrow[m]} \sum_{\substack{i<k: \\ \sigma(i)=\sigma(k)}} 2\alpha_k\alpha_i$$

Performing the sum in the last term over the values of $\sigma(k)$ produces $2\alpha_k s^- \Lambda_{k-1,m}(1)$. It follows,

$$\begin{aligned} \Lambda_{k,m}(x^2, \alpha) &= m\Lambda_{k-1,m}(x^2, \alpha^-) + m\Lambda_{k-1,m-1}(x^2, \alpha^-) \\ &+ m\alpha_k^2\Lambda_{k-1,m}(1) + m\alpha_k^2\Lambda_{k-1,m-1}(1) \\ &+ 2s^-\alpha_k\Lambda_{k-1,m}(1), \end{aligned}$$

and so, when we sum in m to compute Λ_k, the first two rows telescope leaving,

$$\Lambda_k(x^2,\alpha) = 2s^-\alpha_k\Lambda_{k-1}(1).$$

This equals $2\alpha_1\alpha_2$ for $k=2$, and 0 for $k \geq 3$ by Lemma 5. $\square$

Lemma 7. *Assume that each φ_m, $1 \leq m \leq k$ is a (not necessarily symmetric) quadratic polynomial in the α_i. Let $b_m^{(i)}$ $i=0,1,2$ be the sum of the coefficients of the α_j^i terms in φ_m; let $b_m^{(1,1)}$ denote the sum of the coefficients of $\alpha_j\alpha_{j'}$ for $j \neq j'$. If*

$$b_m^\lambda = \begin{cases} b^{(0)}, & \text{if } \lambda = (0), \\ mb^{(i)}, & \text{if } \lambda = (1) \text{ or } \lambda = (2), \\ m(m-1)b^{(1,1)}, & \text{if } \lambda = (1,1), \end{cases}$$

for some b^λ and all m, then

$$\Lambda_k(\varphi,\alpha) = \begin{cases} b^{(0)} + b^{(1)}\alpha_1 + b^{(2)}\alpha_1^2, & \text{if } k=1, \\ 2(b^{(2)} - b^{(1,1)})\alpha_1\alpha_2, & \text{if } k=2, \\ 0, & \text{if } k \geq 3. \end{cases}$$

Proof. Denote by

$$\tilde{\varphi}_m(\alpha) = \frac{1}{n!}\sum_{\sigma:[n]\twoheadrightarrow[n]} \varphi_m(\sigma\alpha),$$

the symmetrized version of φ_m. By definition $\Lambda_{k,m}(\varphi_m,\alpha)$ is symmetric in the α_i. So, for $\sigma : [k] \twoheadrightarrow [m]$ and $y = \sigma\alpha$, if we set $s_\sigma = y_1 + \ldots + y_m$, then $s_\sigma = s = \alpha_1 + \ldots + \alpha_k$ does not depend on σ or m. Thus, $\sum_{i\neq j} y_iy_j = s^2 - \sum_{i=1}^m y_i^2$, and further

$$\tilde{\varphi}_m(y) = b^{(0)} + b^{(1,1)}s^2 + b^{(1)}\sum_{i=1}^m y_i + (b^{(2)} - b^{(1,1)})\sum_{i=1}^m y_i^2.$$

From here we also have

$$\begin{aligned} \Lambda_{k,m}(\varphi_m,\alpha) &= \Lambda_{k,m}(\tilde{\varphi}_m,\alpha) \\ &= (b^{(0)} + b^{(1,1)}s^2)\Lambda_{k,m}(1) + b^{(1)}\Lambda_{k,m}(x,\alpha) + (b^{(2)} - b^{(1,1)})\Lambda_{k,m}(x^2,\alpha), \end{aligned}$$

and the claim follows from Lemmas 5 and 6. $\square$

Lemma 8. *Let now $V = \mathbb{R}^2$, and assume that each $\varphi_m : V^m \to \mathbb{R}$, is a (not necessarily symmetric) quadratic polynomial in $(a_i,\beta_i)_{i=1}^m$. Let b_m^+, b_m^- denote the sum of the coefficients*

of the $\alpha_i\beta_i$ and $\alpha_i\beta_j$, $i \neq j$ terms in φ_m, respectively. Assume as well that all other coefficients vanish, and that there exists b^+ and b^- so that

$$b_m^+ = mb^+, \qquad b_m^- = m(m-1)b^-$$

for all m. Then it holds

$$\Lambda_k(\varphi, \alpha) = \begin{cases} b^+\alpha_1\beta_1, & \text{if } k = 1, \\ (b^+ - b^-)(\alpha_1\beta_2 + \alpha_2\beta_1), & \text{if } k = 2, \\ 0 & \text{if }, k \geq 3. \end{cases}$$

Proof. By Lemma 7 we may assume that each φ_m is symmetric. Setting $s_\beta = \beta_1 + \ldots + \beta_k$, $s_\alpha = \alpha_1 + \ldots + \alpha_k$, and for $\sigma : [m] \twoheadrightarrow [k]$ setting $y = \sigma\alpha$, $w = \sigma\beta$, the arguments of Lemma 7 give

$$\varphi_m((y_1, w_1), \ldots, (y_m, w_m)) = b^- s_\alpha s_\beta + (b^+ - b^-) \sum_{i=1}^m y_i w_i.$$

We use polarization ($4y_iw_i = (y_i + w_i)^2 - (y_i - w_i)^2$) for the second term and find that,

$$\Lambda_{k,m}(\phi, (\alpha, \beta)) = b^- s_\alpha s_\beta \Lambda_{k,m}(1) + \frac{b^+ - b^-}{4} \left(\Lambda_{k,m}(x^2, \alpha + \beta) - \Lambda_{k,m}(x^2, \alpha - \beta)\right).$$

The proof is finished by invoking Lemmas 5 and 6. □

Lemma 9 (Spohn, Soshnikov). *Let $\varphi_m(\gamma_1, \ldots, \gamma_m) = \max(\eta_1, \ldots, \eta_m)$, where $\eta_j = \gamma_1 + \ldots + \gamma_j$. If $\alpha_1 + \ldots + \alpha_k = 0$, then*

$$\Lambda_k(\varphi, \alpha) = \begin{cases} -|\alpha_1| = -|\alpha_2|, & \text{if } k = 2, \\ 0, & \text{if } k \neq 2. \end{cases}$$

A result of this type was first discussed by Spohn in [21] (though see [20] as well) which focusses on the determinantal process on the line with the sine kernel $\frac{\sin(\pi(x-y))}{\pi(x-y)}$. It also appears as the 'Main Combinatorial Lemma' in the the paper of Soshnikov, [23], where it is used to track fluctuations of the eigenvalues of the Unitary group $U(n)$.

In our result, Lemma 9 figures into the boundary component of the limiting noise, and so may be thought of as the root of the connection to $U(n)$.

We finish with the note that in the formulation of [23], the $\eta'_j = \gamma'_1 + \ldots + \gamma'_{j-1}$. That is, the sum only goes up to index $j-1$). This is equivalent to the above upon setting $\gamma'_i = -\gamma_{k+1-i}$, so that $\eta'_j = \gamma'_1 + \ldots + \gamma'_{j-1} = -(\gamma_k + \ldots + \gamma_{k+1-(j-1)}) = \eta_{k+1-j}$ and $\max(\eta'_1, \ldots, \eta'_m) = \max(\eta_1, \ldots, \eta_m)$.

6 Polynomial statistics for the Ginibre ensemble

We now apply the results of the previous two sections to the Ginibre ensemble. That is, in (8) we put $L = n < \infty$, $\mu(z) = \mu_n(z) = \frac{n}{\pi} e^{-n|z|^2}$, $c_k = \frac{n^k}{k!}$, $\lambda_k = 1$, and then take $n \to \infty$. Further, while the possibility of using weighted polynomials may be important in general, in the case of Ginibre it suffices to work with the naked monomials $\zeta(z) = z$.

With these specifications,

$$M(2\ell + 2\eta_j + \alpha_j - \beta_j, \alpha_j + \beta_j) = M(2\ell + 2\eta_j + 2\alpha_j),$$

and plainly

$$\frac{M(2q + 2\kappa)}{M(2\kappa)} = \frac{n^q (q+\kappa)!}{q! n^{q+\kappa}} = n^{-\kappa}(q+1)\cdots(q+\kappa).$$

Substituting into (10), we have

$$\Phi_m(z_1^{\alpha_1} \bar{z}_1^{\beta_1}, \ldots, z_m^{\alpha_m} \bar{z}_m^{\beta_m}) = \frac{1}{n^s} \sum_{\ell=-\eta_{min}}^{n-1-\eta_{max}} \prod_{j=1}^{m} (\ell + \eta_j + 1) \cdots (\ell + \eta_j + \alpha_j), \tag{11}$$

which for large n reads,

$$\frac{1}{n^s} \sum_{\ell=-\eta_{min}}^{n-1-\eta_{max}} \left[\ell^s + \ell^{s-1} \sum_{j=1}^{m} \left[\eta_j \alpha_j + \binom{\alpha_j + 1}{2} \right] + O(\ell^{s-2}) \right],$$

where again $s = \sum \alpha_i = \sum \beta_i$. Next, use the fact that

$$\sum_{\ell = \text{const.}}^{n} \ell^s = \frac{n^{s+1}}{s+1} + \frac{n^s}{2} + O(n^{s-1}),$$

valid for any $s \geq 1$, to put the the asymptotics of (11) into the form:

$$\Phi_m(f_1, \ldots, f_m) = \frac{n}{s+1} - (1 + \eta_{\max}) + \frac{1}{2} + \frac{1}{s} \sum_{j=1}^{m} \left[\eta_j \alpha_j + \binom{\alpha_j + 1}{2} \right] + O(n^{-1}). \tag{12}$$

Here $f_\ell = f_\ell(z, \bar{z})$ is shorthand for $z^{\alpha_\ell} \bar{z}^{\beta_\ell}$.

For fixed s, the $O(1)$ term in (12) is a function of the α_i and β_i, which we denote φ_m. Summing the above into the basic cumulant term, we find that

$$\Upsilon_k(\Phi, (f_1, \ldots, f_k)) = \frac{n}{s+1} \Lambda_k(1) + \Lambda_k(\varphi, (\alpha, \beta)) + O(n^{-1}), \tag{13}$$

and are in position to use the combinatorial tools developed in Section 5.

First of all, the coefficient of n in (13) vanishes for $k \geq 2$ by Lemma 5. For the constant order term, examining φ_m shows it to be a linear combination of η_{max}, a quadratic polynomial

in α_i and a polynomial which is itself a linear combination of the $\alpha_i\beta_j$. Further, the sums of the coefficients in the polynomial part are as follows:

1	α_i	α_i^2	$\alpha_i\alpha_j, i \neq j$	$\alpha_i\beta_i$	$\alpha_i\beta_j, i \neq j$
$-1/2$	$m/(2s)$	$-m/(2s)$	$-m(m-1)/(2s)$	m/s	$m(m-1)/(2s)$

Now combined, Lemmas 7, 8 and 9 imply that $\Lambda_k(\varphi, (\alpha, \beta))$, and so the full cumulant, vanishes in the limit for $k > 2$. In summary we have:

Theorem 10. *For any real-valued polynomial $P(z, \bar{z}) = \sum c_{\alpha_k,\beta_k} z^{\alpha_k} \bar{z}^{\beta_k}$, the Ginibre statistic $X_n(P) - \mathbf{E}[X_n(P)]$ converges in distribution to a mean-zero Gaussian as $n \to \infty$.*

As for the limiting variance, choose a pair of monomials $z^{\alpha_1}\bar{z}^{\beta_1}$ and $z^{\alpha_2}\bar{z}^{\beta_2}$ with $\alpha_1+\alpha_2 = \beta_1 + \beta_2$, and note that our cumulant asymptotics enter the picture as in

$$\begin{aligned} \lim_{n\to\infty} \mathrm{Cov}(z^{\alpha_1}\bar{z}^{\beta_1}, z^{\beta_2}\bar{z}^{\alpha_2}) &= \lim_{n\to\infty} \Big(\Phi_1(z^{\alpha_1+\alpha_2}\bar{z}^{\beta_1+\beta_2}) - \Phi_2(z^{\alpha_1}\bar{z}^{\beta_1}, z^{\alpha_2}\bar{z}^{\beta_2}) \Big) \\ &= \max(0, \beta_1 - \alpha_1) + \frac{\alpha_1\beta_2}{s}. \end{aligned} \tag{14}$$

On the left, $z^{\alpha_1}\bar{z}^{\beta_1}$ and $z^{\beta_2}\bar{z}^{\alpha_2}$ stand in for the full linear statistics, and α_2, β_2 switch rolls from left to right since covariance is conjugate linear in the second argument. The limit itself is read off from (12); why the corresponding covariance vanishes in the limit if $\alpha_1+\alpha_2 \neq \beta_1+\beta_2$ should also be clear.

Next observe that

$$\frac{1}{\pi}\int_{\mathbb{U}} \bar{\partial}(z^{\alpha_1}\bar{z}^{\beta_1})\overline{\bar{\partial}(z^{\beta_2}\bar{z}^{\alpha_2})}d^2z = \frac{\alpha_2\beta_1}{\alpha_1+\alpha_2} = \frac{\alpha_2\beta_1}{s}, \tag{15}$$

and

$$\sum_{k>0} k\,(z^{\alpha_1}\bar{z}^{\beta_1})\hat{}(k)\,\overline{(z^{\beta_2}\bar{z}^{\alpha_2})\hat{}(k)} = \max(0, \alpha_1 - \beta_1). \tag{16}$$

A little algebra will show that the sum of (15) and (16) equals the final expression in (14). By linearity we may conclude: with any real-valued polynomials f and g in z and $\bar{z}$,

$$\lim_{n\to\infty} \mathrm{Cov}(f, g) = \frac{1}{4\pi}\langle f, g\rangle_{H^1(\mathbb{U})} + \frac{1}{2}\langle f, g\rangle_{H^{1/2}(\partial\mathbb{U})}.$$

The general covariance/variance asymptotics occupy the next section.

7 Concentration

This section is devoted to the following estimate.

Theorem 11. *For linear statistics $X_n(f)$ and $X_n(g)$ in the Ginibre ensemble,*

$$\lim_{n\to\infty} \mathrm{Cov}(X_n(f), X_n(g)) = \frac{1}{\pi}\int_{\mathbb{U}} \bar{\partial} f(z)\overline{\bar{\partial} g}(z)\, d^2 z + \sum_{k>0} k\widehat{f}(k)\,\overline{\widehat{g}(k)}, \tag{17}$$

as long as f and g possess continuous partial derivatives in a neighborhood of $\mathbb{U}$ and are otherwise bounded as in $|f(z)| \vee |g(z)| \le Ce^{c|z|}$.

Granting (17), the central limit theorem extends immediately from polynomials of the form

$$P_m(z,\bar{z}) = \sum_{\alpha_k+\beta_k\le M} c_{\alpha_k,\beta_k} z^{\alpha_k}\bar{z}^{\beta_k},$$

to general test functions $f(z)$ satisfying the growth and regularity conditions assumed in the statement.

First, by the Stone-Weirstrauss theorem, we can find a polynomial $P'(z,\bar{z})$ such that $\|P'(z,\bar{z}) - \bar{d}f\|_{L^\infty(|z|\le 1+\delta)} \le \varepsilon$ for whatever $\varepsilon > 0$. Certainly the L^∞-norm (of the derivatives) on the larger disk controls both the $H^1(\mathbb{U})$ and $H^{1/2}(\partial\mathbb{U})$ norms, and there exists a sequence of polynomials $P_m(z,\bar{z})$ such that

$$\Big|\Big| f - P_m \Big|\Big|_{H^1(\mathbb{U})\cap H^{1/2}(\mathbb{U})} \to 0, \tag{18}$$

as the degree $m\to\infty$. At each step, P_m is just the anti-derivative (in $\bar{z}$) of the polynomial approximating $\bar{\partial} f$.

Next, denote the centered statistic by

$$Z_n(f) = X_n(f) - \frac{n}{\pi}\int_{\mathbb{U}} f(z) d^2 z.$$

With f as in the above result, a separate (and easy) calculation shows that

$$\mathbf{E}[X_n(f)] - \frac{n}{\pi}\int_{\mathbb{U}} f(z) d^2 z = o(1),$$

and it follows that the family $\{Z_n(f)\}$ is tight as $n\to\infty$. Denote a given subsequential limit by $Z_\infty(f)$. For polynomial test functions, we have proved that $Z_n(P_m)$ converges in distribution to a mean zero complex Gaussian, $Z_\infty(P_m)$. Now pick a smooth bounded $\phi : \mathbb{C}\to[0,\infty)$ with $\|\nabla\phi\|_\infty \le 1$, and note that,

$$\begin{aligned}\Big|\mathbf{E}\phi(Z_\infty(f)) - \mathbf{E}\phi(Z_\infty(P_m))\Big| &\le \limsup_{N\to\infty} \mathbf{E}\Big|\phi(Z_n(f)) - \phi(Z_n(P_m))\Big| \\ &\le \limsup_{n\to\infty} \mathbf{E}[|Z_n(f-P_m)|^2]^{1/2},\end{aligned}$$

where things have been fixed so the right hand side tends to zero as $m\to\infty$. This appraisal is independent of the special subsequence chosen in the definition of $Z_\infty(f)$. Hence, by appealing once more to Theorem 11 and (18) the large m limit of $Z_\infty(P_m)$ identifies the limit distribution of $Z_n(f)$ unambiguously and completes the proof of Theorem 1.

Verification of the limiting variance

The opening move in the proof of Theorem 11 is the covariance formula,

$$\begin{aligned}\mathrm{Cov}\big(X_n(f), X_n(g)\big) &= \int_{\mathbb{C}} f(z)\,\overline{g(z)}\,K_n(z,\bar{z})\,d\mu_n(z) \\ &\quad - \int_{\mathbb{C}}\int_{\mathbb{C}} f(z)\,\overline{g(w)}\,|K_n(z,\bar{w})|^2 d\mu_n(z)\,d\mu_n(w),\end{aligned}$$

valid for any determinantal point process. Then, each appearance of the test functions f and g is expanded by way of the dbar representation. That is the fact that, for any f once continuously differentiable in a domain $\mathbb{D} \subset \mathbb{C}$ and continuous up to the boundary $\partial\mathbb{D}$, it holds,

$$f(\zeta) = -\frac{1}{\pi i}\int_{\mathbb{D}} \frac{\bar{\partial} f(w)}{w-\zeta}\,d^2w \;+\; \frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{f(w)}{w-\zeta}\,dw, \tag{19}$$

see for example [4].

The formula (19) allows one to pass from the general variance asymptotics to one for the particular functional $z \mapsto (z-\cdot)^{-1}$. The decay properties of the Ginibre ensemble allow further simplifications. In particular, with $f(z)$ once differentiable in say $|z| < 1+\varepsilon$, decompose as in $f(z) = f(z)\psi(z) + f(z)(1-\psi(z))$ for a smooth ψ, equal to 1 on $\mathbb{U}$ and vanishing for $|z| > 1+\varepsilon/2$. Next, for any $\varepsilon > 0$ and $c > 0$,

$$\begin{aligned}\int_{|z|\geq 1+\varepsilon} e^{c|z|^2} K_n(z,\bar{z}) d\mu_N(z) &= \sum_{\ell=0}^{n-1} \frac{n^{\ell+1}}{\pi \ell!}\int_{|z|\geq 1+\varepsilon} e^{c|z|^2}|z|^{2\ell}e^{-n|z|^2}d^2z \\ &\leq n \times \frac{n^n}{(n-1)!}\int_{(1+\varepsilon)^2}^{\infty} r^{n-1}e^{-(n-c)r}\,dr \leq C' n e^{-n\varepsilon^2/4},\end{aligned} \tag{20}$$

by a simple Laplace estimate. From the growth assumption on f and noting $|K(z,\bar{w})|^2 \leq K(z,\bar{z})K(w,\bar{w})$, it follows that the variance of $X_n(f)$ agrees with that of $X_n(f\psi)$ up to $o(1)$ errors. The same pertains to the covariance, and it suffices to take the test functions compactly supported from the start, ignoring the boundary integral in the dbar representation (19). In short, the covariance may be read off from the $n \to \infty$ behavior of

$$\begin{aligned}&\frac{1}{\pi^2}\int_{|\nu|<1+\varepsilon}\int_{|\eta|<1+\varepsilon} \bar{\partial} f(\nu)\overline{\bar{\partial} g(\eta)} \times \\ &\quad \left\{ \int_{\mathbb{C}} \frac{1}{z-\nu}\,\overline{\frac{1}{z-\eta}}\,K_n(z,\bar{z})\,d\mu(z) - \int_{\mathbb{C}}\int_{\mathbb{C}} \frac{1}{z-\nu}\,\overline{\frac{1}{w-\eta}}|K_n(z,\bar{w})|^2\,d\mu_n(z)d\mu_n(w)\right\} d^2\nu d^2\eta,\end{aligned} \tag{21}$$

in which $f(z)$ and $g(z)$ and their first partial derivatives are assumed to vanish along $|z| = 1+\varepsilon$.

Defining,

$$\begin{aligned}\phi_n(\nu,\eta) \;&=\; \int_{\mathbb{C}} \frac{1}{z-\nu}\,\overline{\frac{1}{z-\eta}}\,K_n(z,\bar z)\,d\mu_n(z) \\ &\quad -\int_{\mathbb{C}}\int_{\mathbb{C}} \frac{1}{z-\nu}\,\overline{\frac{1}{w-\eta}}\,K_n(z,\bar w)K_n(w,\bar z)\,d\mu_n(z)\,d\mu_n(w),\end{aligned} \tag{22}$$

the proof of Theorem 11 splits into two parts. For the interior (or H^1) contribution we have:

Lemma 12. *For fixed ν with $|\nu|<1$,*

$$\phi_n(\nu,\bar\eta) \to \pi\,\delta_\nu(\eta),$$

in the sense of measures as $n\to\infty$. It follows that,

$$\lim_{N\to\infty}\frac{1}{\pi^2}\int_{|\nu|<1}\int_{|\eta|<1+\varepsilon} \bar\partial f(\nu)\overline{\bar\partial f(\eta)}\phi_N(\nu,\bar\eta)\,d^2\nu d^2\eta = \frac{1}{\pi}\int_{\mathbb{U}} |\bar\partial f(\nu)|^2 d^2\nu,$$

for bounded and continuous $\bar\partial f$.

For the boundary, (or $H^{1/2}$) contribution, we prove separately:

Lemma 13. *With $H^{1/2}_+(\partial\mathbb{U})$ corresponding to the usual sum extended only over positive indices, it holds that*

$$\lim_{n\to\infty}\frac{1}{\pi^2}\int_{1<|\nu|,|\eta|<1+\varepsilon} \bar\partial f(\nu)\overline{\bar\partial f(\eta)}\,\phi_n(\nu,\bar\eta)\,d^2\nu\,d^2\eta = \|f\|^2_{H^{1/2}_+(\partial\mathbb{U})},$$

whenever f is continuously differentiable and vanishes on $|z|=1+\varepsilon$.

Note that we have set $f=g$ in the above two statements. This has only been done for the sake of slightly more compact expressions. No generality is lost: covariances can always be recovered from variances.

Proof of Lemma 12. We begin by showing that,

$$\lim_{n\to\infty}\int_{\mathbb{B}} \phi_n(\nu,\bar\eta)\,d^2\eta = \pi, \tag{23}$$

for $|\nu|<1$ and $\mathbb{B}$ any disk about the origin containing ν. A residue calculation gives,

$$\int_{|\eta|<b}\frac{d^2\eta}{z-\eta} = \begin{cases} \pi\bar z & \text{for } |z|<b \\ \frac{\pi b^2}{z} & \text{for } |z|>b \end{cases},$$

and so, denoting the radius of $\mathbb{B}$ by $b > |\nu| = a$,

$$\begin{aligned}\int_{\mathbb{B}} \phi_n(\nu, \bar\eta)\, d^2\eta \;&=\; \pi \int_{a<|z|<b} K_n(z, \bar z)\, d\mu_n(z) + \pi b^2 \int_{|z|>b} \frac{1}{|z|^2} K_n(z, \bar z)\, d\mu_n(z) \qquad (24)\\ &\quad -\pi \sum_{\ell=0}^{n-2} c_\ell c_{\ell+1} \int_{|z|>a} |z|^{2\ell}\, d\mu_n(z) \int_{|w|<b} |w|^{2\ell+2}\, d\mu_n(w)\\ &\quad -\pi b^2 \sum_{\ell=0}^{n-2} c_\ell c_{\ell+1} \int_{|z|>a} |z|^{2\ell}\, d\mu_n(z) \int_{|w|>b} |w|^{2\ell}\, d\mu_n(w),\end{aligned}$$

after expanding $(z-\nu)^{-1}$ in series where $|z| > |\nu|$. Recall that $c_k = c_k^n = \frac{n^k}{k!}$. Now recombine (24) in the form

$$\begin{aligned}\int_{\mathbb{B}} \phi_n(\nu, \eta) d^2\eta \;&=\; \pi c_{N-1} \int_{|z|>a} |z|^{2(n-1)}\, d\mu_n(z) - \pi \int_{|z|>b} \Big(1 - \frac{b^2}{|z|^2}\Big)\, d\mu_n(z) \qquad (25)\\ &\quad - \pi \sum_{\ell=0}^{n-2} c_\ell c_{\ell+1} \int_{|z|<a} \int_{|w|>b} |z|^{2\ell} |w|^{2\ell+2} \Big(1 - \frac{b^2}{|w|^2}\Big)\, d\mu_n(z)\, d\mu_n(w).\end{aligned}$$

A simple computation shows that the first two terms of (25) tend to π and zero respectively as $n \to \infty$. The final two terms of (25) may in turn be bounded by a constant multiple of

$$\sum_{\ell=0}^{n-2} c_\ell c_{\ell+1} \int_{|z|<a} \int_{|w|>b} |z|^{2\ell} |w|^{2\ell+2}\, d\mu_n(z)\, d\mu_n(w) \le \sum_{\ell=0}^{n-2} P\Big(S_\ell^n \le a^2\Big) P\Big(S_{\ell+1}^n \ge b^2\Big) \qquad (26)$$

where S_ℓ^n denotes a sum of $(\ell+1)$ independent mean $1/n$ exponential random variables. Standard large deviation estimates explain why $P(S_\ell^n < a^2)$ is exponentially small in n if $\ell > n(a^2 + \delta)$, with an estimate of the same type holding for $P(S_\ell^n > b^2)$ and $\ell < n(b^2 - \delta)$. Since there is a gap, *i.e.* $a < b$, each term in the above sum is of order $e^{-\gamma n}$ for a γ positive with $(b-a)^2$. Since there are only n terms, the verification of (23) is complete.

It remains to show that we have decay away from $\eta = \nu$. Given $\delta > 0$, we claim that

$$\lim_{n\to\infty} \int_{\mathbb{A}\cap|\nu-\eta|>\delta} \Big|\phi_n(\nu, \bar\eta)\Big|\, d^2\eta = 0,$$

with any $\mathbb{A}$ supported in $|z| \le 1 + \varepsilon$.

First fix η as well as ν, and take $a = |\nu| \le |\eta| = b$ without any loss. Two types of expressions emerge from performing the integrations over z and w in the definition of $\phi_n(\nu, \bar\eta)$, an inner term corresponding to $|z| < a$ and $|w| < b$, and an outer term corresponding to $|z| > a$ and $|w| > b$:

$$\phi_n(\nu, \bar\eta) = \phi_o(\nu, \bar\eta) + \phi_e(\nu, \bar\eta).$$

The integration over the mixed regions ($|z| < a$, $|w| > b$, for example) vanishes by orthogonality. The inner term reads

$$\begin{aligned}
\phi_o(\nu,\bar{\eta}) &= \sum_{\ell=1}^{n-1} \frac{1}{(\nu\bar{\eta})^{\ell+1}} \sum_{m=n-\ell}^{n-1} c_m \int_{|z|<a} |z|^{2(m+\ell)} d\mu_n(z) \\
&\quad + \sum_{\ell=n}^{\infty} \frac{1}{(\nu\bar{\eta})^{\ell+1}} \sum_{m=0}^{n-1} c_m \int_{|z|<a} |z|^{2(m+\ell)} d\mu_n(z) \\
&\quad + \sum_{\ell=0}^{n-1} \frac{1}{(\nu\bar{\eta})^{\ell+1}} \sum_{m=0}^{n-\ell-1} c_m c_{m+\ell} \int_{|z|<a} |z|^{2(m+\ell)} d\mu_n(z) \int_{|z|>b} |z|^{2(m+\ell)} d\mu_n(z) \\
&= \phi_o^{(1)}(\nu,\bar{\eta}) + \phi_o^{(2)}(\nu,\bar{\eta}) + \phi_o^{(3)}(\nu,\bar{\eta}).
\end{aligned} \tag{27}$$

The result of the outer integration has the same shape, up to the obvious inversions. For example, the analog of the first term on the right of (27) is $\phi_e^{(1)} = \sum_{\ell=1}^{n-1} \sum_{m=n-\ell}^{n-1} (\nu\bar{\eta})^{\ell}$ $c_m \int_{|z|>b} |z|^{2(m-\ell-1)} d\mu_n(z)$. We sketch the proof of the L^1 decay of ϕ_o. The estimates for ϕ_e are much the same, and both cases are similar to considerations immediately above.

Integrating by parts produces the bounds,

$$\frac{n}{k+1}\, a^{2k+2}\, e^{-na^2} \le \int_{|z|<a} |z|^{2k} d\mu_n(z) \le \frac{1}{1-\frac{n}{k}a^2}\, \frac{n}{k+1}\, a^{2k+2}\, e^{-na^2}, \tag{28}$$

for $k > n$. Recall here that $a < 1$. If we further assume that $b > a$, there is a constant C such that

$$\begin{aligned}
\left|\phi_o^{(1)}(\nu,\bar{\eta})\right| &\le C\, e^{-na^2} \sum_{\ell=0}^{n-1} \sum_{m=0}^{\ell-1} \frac{n^{n-(\ell-m)}}{(n-(\ell-m))!} a^{2(n-(\ell-m))}\, (a/b)^{\ell} \\
&\le C\, e^{-na^2} \sum_{k=0}^{n-1} \frac{(na^2)^k}{k!} (a/b)^{n-k},
\end{aligned}$$

(after changing variables and the order of summation) and

$$\left|\phi_o^{(2)}(\nu,\bar{\eta})\right| \le C \sum_{\ell=n}^{\infty} (a/b)^{\ell} \Big(e^{-na^2} \sum_{k=0}^{n-1} \frac{(na^2)^k}{k!} \Big) \le C \frac{(a/b)^n}{(1-a/b)}.$$

The last two displays clearly tend to zero as $n \to \infty$, providing the advertised L^1-decay over sets supported away from $|\nu| = |\eta|$. To finish it is enough that both $\phi_o^{(1)}(\nu,\bar{\eta})$ and $\phi_o^{(2)}(\nu,\bar{\eta})$ remain bounded along that part of the circle $|\nu| = |\eta|$ where $\nu/\eta = e^{i\theta}$ and say $|\theta| > \delta/2$. This is a simple exercise taking advantage of the oscillations introduced in summing powers of $e^{i\theta}$

As for $\phi_o^{(3)}(\nu,\bar\eta)$, the estimate (28) does not have the same effect as $m+\ell \le n$ in each appearance of $\int |z|^{2(m+\ell)} d\mu_n(z)$. Instead, one works along the lines of (26). Again, first take $a<b$ and write,

$$\left|\phi_o^{(3)}(\nu,\bar\eta)\right| \le \sum_{\ell=1}^{n-1} (ab)^{-\ell} \sum_{m=0}^{n-\ell-1} \frac{c_{m+\ell}}{c_m} P(S_{m+\ell}^n \le a^2) P(S_{m+\ell}^n \ge b^2).$$

Focussing on the sum restricted to $\ell + m \le na^2$, that object is bounded by

$$\begin{aligned} P(S_{na^2} \ge b^2) \sum_{\ell=0}^{na^2} (ab)^{-\ell} \sum_{m=0}^{na^2-\ell} \frac{c_{m+\ell}}{c_m} &\le 2e^{-n(b-a)^2/4} \sum_{\ell=0}^{na^2} (ab)^{-\ell} \sum_{m=0}^{na^2-\ell} \frac{(m^\ell + \ell^2 m^{\ell-1})}{n^\ell} \\ &\le Cne^{-n(b-a)^2/4} \sum_{\ell=0}^{na^2} \ell (a/b)^\ell \to 0, \end{aligned}$$

having once more used the standard large deviations estimate for exponential random variables. The sums over $m+\ell \ge nb^2$ and $na^2 \le m+\ell \le nb^2$ are handled by the same procedure. When $a=b$ but ν/η is kept away from 1, the key observation are that $P(S_{m+\ell}^n \le a^2) P(S_{m+\ell}^n \ge a^2)$ has a good decay away from $|m+\ell-na^2| \le C\sqrt{n}$. By using this in conjunction with the oscillations from $(\nu/\eta)^{-\ell}$, the boundedness of $\phi_o^{(3)}(\nu,\bar\eta)$ along $\{|\nu|=|\eta|\} \cap \{|\nu-\eta|>\delta\}$ will follow. $\square$

Proof of Lemma 13. Start with the point-wise limit of $\phi_n(\nu,\bar\eta)$ for which we fix $\nu \ne \eta$ in the annulus $1 < |\cdot| < 1+\varepsilon$. For the interior estimate, it was be convenient to carry the integration in z and w in the definition of ϕ_n over all of $\mathbb{C}$, see (22). However, by the estimate (20), one may cut down the integral to $|\cdot| < 1+\delta$ for any $\delta > 0$ at the expense of an exponentially small error. We therefore consider the $n \to \infty$ limit of

$$\begin{aligned} \widetilde{\phi}_n(\nu,\bar\eta) &= \int_{|z|<1+\delta} \frac{1}{(z-\nu)(\overline{z-\eta})} K_n(z,\bar z) d\mu_n(z) \\ &\quad - \int_{|z|,|w|<1+\delta} \frac{1}{(z-\nu)(\overline{w-\eta})} |K_n(z,\bar w)|^2 \, d\mu_n(z) d\mu_n(w), \end{aligned}$$

in which δ is chosen so that $1+\delta$ is less than either $|\nu|$ or $|\eta|$.

Next, each appearance of $(\cdot-\eta)^{-1}$ and $(\cdot-\nu)^{-1}$ in $\widetilde{\phi}_n$ is expanded in series, and the

integrals are performed term-wise to find that,

$$\begin{aligned}
\widetilde{\phi}_n(\nu,\bar{\eta}) &= \sum_{\ell=0}^{\infty}\int_{|z|<1+\delta}\frac{|z|^{2\ell}}{(\nu\bar{\eta})^{\ell+1}}K_n(z,\bar{z})d\mu_n(z)\\
&\quad -\sum_{\ell=0}^{n-1}\int_{|z|<1+\delta}\int_{|w|<1+\delta}\frac{(z\bar{w})^k}{(\nu\bar{\eta})^{\ell+1}}|K_n(z,\bar{w})|^2 d\mu_n(z)d\mu_n(w)\\
&= \sum_{\ell=0}^{\infty}\frac{1}{(\nu\bar{\eta})^{\ell+1}}\sum_{m=0}^{n-1}c_m\int_{|z|<1+\delta}|z|^{2(m+\ell)}d\mu_n(z)\\
&\quad -\sum_{\ell=0}^{n-1}\sum_{m=0}^{n-\ell-1}c_m c_{m+\ell}\Big(\int_{|z|<1+\delta}|z|^{2(m+\ell)}d\mu_n(z)\Big)^2.
\end{aligned}$$

In the first equality, only the diagonal terms survive in the expansion an account of orthogonality. Now note that, for any $\alpha>0$,

$$\sum_{\alpha n\le\ell\le\infty}\int_{|z|<1+\delta}\frac{|z|^{2\ell}}{|\nu\eta|^{\ell+1}}K_n(z,\bar{z})\mu_n(dz)\le n\sum_{\ell=\lfloor\alpha n\rfloor}^{\infty}(1-\theta_\delta)^\ell = n\theta_\delta^{-1}(1-\theta_\delta)^{\alpha n},$$

in which $\theta_\delta>0$ for $\delta>0$, and

$$1-c_k\int_{|z|<1+\delta}|z|^{2k}d\mu_n(z)=O(e^{-n(\delta-\alpha)^2})$$

for any $k\le(1+\alpha)n$ with $\alpha<\delta$. It follows that,

$$\widetilde{\phi}_n(\nu,\bar{\eta})=\sum_{\ell=1}^{\lceil\alpha n\rceil}\frac{1}{(\nu\bar{\eta})^{\ell+1}}\sum_{m=n-\ell}^{n-1}\frac{c_m}{c_{m+\ell}}+o(1),$$

with α obeying both the above constraints. Finally, since

$$\sum_{m=n-\ell}^{n-1}\frac{c_m}{c_{m+\ell}}=\sum_{m=0}^{\ell-1}\Big(1+\frac{m}{n}\Big)\Big(1+\frac{m-1}{n}\Big)\cdots\Big(1+\frac{m-(\ell-1)}{n}\Big)=\ell+O\Big(\frac{\ell^3}{n^2}\Big)$$

we have that

$$\lim_{n\to\infty}\phi_n(\nu,\bar{\eta})=\sum_{\ell=1}^{\infty}\frac{\ell}{(\nu\bar{\eta})^{\ell+1}}.$$

The limit holds uniformly on $\{|\nu|\ge 1+\delta\cup|\eta|\ge 1+\delta\}$ for any $\delta>0$, and a dominated convergence argument produces

$$\lim_{n\to\infty}\frac{1}{\pi^2}\int_{1<|\nu|,|\eta|<1+\varepsilon}\bar{\partial}f(\nu)\overline{\bar{\partial}f(\eta)}\phi_n(\nu,\bar{\eta})\,d^2\nu d^2\eta=\sum_{\ell=1}^{\infty}\ell\,\Big|\frac{1}{\pi}\int_{1<|\eta|<1+\varepsilon}\frac{\bar{\partial}f(\eta)}{\eta^{\ell+1}}\,d^2\eta\Big|^2.$$

It remains to realize that since $\bar{\partial}[f(\eta)\eta^{-k}(\eta-\omega)] = \eta^{-k}(\eta-\omega)(\bar{\partial}f)(\eta)$ for η away from zero and any ω, the dbar formula reads,

$$0 = -\frac{1}{\pi}\int_{1<|\eta|<1+\varepsilon} \frac{\bar{\partial}f(\eta)}{\eta^{\ell+1}}\, d^2\eta + \frac{1}{2\pi}\int_{|\eta|=1} \frac{f(\eta)}{\eta^{\ell+1}}\, d\eta + \frac{1}{2\pi}\int_{|\eta|=1+\varepsilon} \frac{f(\eta)}{\eta^{\ell+1}}\, d\eta.$$

The third term vanishes by assumption, and the second term is precisely $\sqrt{-1}$ times the ℓ-the Fourier coefficient of the boundary data of f. $\square$

8 Universality for analytic functionals

For each positive integer n, pick a rotation invariant measure, $d\mu_n(z) = \tilde{\mu}_n(|z|)d^2z$, and set

$$K_n(z,\bar{w}) = \sum_{\ell=0}^{n-1} c_{n,\ell}(z\bar{w})^\ell, \quad \text{where} \quad c_{n,\ell}^{-1} = M(n,2\ell) = \int_{\mathbb{C}} |z|^{2\ell} d\mu_n(z);$$

the measure μ_n itself need not depend explicitly on n. Now impose the following moment condition: for all integer m there is a $\rho > 0$ such that

$$M(n,2n+2m)/M(n,2n) \to \rho^{2m}, \tag{29}$$

as $n \to \infty$. Of course, by a scaling, it may be assumed that $\rho = 1$.

For any determinantal point process (K,μ) in $\mathbb{C}$ with rotation invariant μ, it is the case that the moduli $\{|z_1|, |z_2|, \ldots, |z_n|\}$ as a set have the same distribution as $\{R_1, \ldots, R_k\}$ where R_i are independent with $P(R_k \in dr) = 2\pi c_k r^{2k+1}\ \tilde{\mu}(r)dr$ for each k (see [11] or [14]). Therefore, (29) has the interpretation that the stochastically largest random variable R_n satisfies $\mathbf{E}_n R_n^{2m} \to \rho$, for every $m \in \mathbb{Z}$. Equivalently, for every fixed $\ell \geq 0$ we have $\mathbf{E}_n R_{n-\ell}^2 \to \rho$.

Being the main example at hand, the reader will be happy to check that Ginibre satisfies everything asked for: the relevant computation is $\lim_{n\to\infty} \frac{\Gamma(n+m)}{n^m\Gamma(n)} = 1$. A second example of interest is the *truncated Bergman ensemble*. Here one begins with $\mu_n =$ the uniform measure on $\mathbb{U}$, producing the kernel

$$K_n(z,\bar{w}) = \sum_{\ell=0}^{n-1}(\ell+1)z^\ell\bar{w}^\ell$$

of orthonormal polynomials on the disk. This model arises naturally in the following way. Consider the random polynomial $z^n + \sum_{k=0}^{n-1} a_k z^k$ with independent coefficients drawn uniformly from the large disk $R\,\mathbb{U}$. If we condition the roots to lie entirely within the unit disk, the $R \to \infty$ limit of the resulting point process is the truncated Bergman ensemble. This

observation may be gleaned from Hammersely [10], but see also [17] for why the adjective "truncated" is used.

Consider now a linear statistic in any such ensemble which is polynomial in z alone, or of the form $\sum_{i=1}^{n} p_m(z_i) = \sum_{k=1}^{m} a_k \Big(\sum_{i=1}^{n} z_i^k\Big)$. We have the following CLT.

Theorem 14. *Let $z_1, \dots, z_n$ be drawn from a determinantal process (K_n, μ_n) as above for which (29) holds with $\rho = 1$. Denote $p_j(z^{\oplus n}) = z_1^j + \cdots + z_n^j$. Take $a = (a_1, \dots, a_k)$ and $b = (b_1, \dots, b_k)$ with $a_j, b_j \in \{0, 1, \dots\}$ and bounded independently of n. Then,*

$$\lim_{n\to\infty} \mathbf{E}_n \left[\prod_{j=1}^{k} p_j(z^{\oplus n})^{a_j} \overline{p_j(z^{\oplus n})}^{b_j} \right] = \mathbf{E} \left[\prod_{j=1}^{k} \Big(\sqrt{j} Z_j\Big)^{a_j} \overline{\Big(\sqrt{j} Z_j\Big)}^{b_j} \right], \tag{30}$$

for $Z_1, \dots, Z_k$ independent standard complex Gaussian random variables.

Having identified the limiting moments, this implies Theorem 3. On a case by case basis, Theorem 14 and a variance estimate yields a more complete picture.

Corollary 15. *Let $f(z)$ be analytic in a neighborhood of $|z| \le 1$ and the points $z_1, \dots, z_n$ be drawn either from the Ginibre or truncated Bergman ensemble. Then, as $n \to \infty$, $\sum_{\ell=1}^{n} f(z_\ell) - n f(0)$ converges in distribution to a mean-zero complex Normal with variance $\frac{1}{\pi} \int_{\mathbb{U}} |f'(z)|^2 \, d^2 z$.*

Qualitatively, the Ginibre and truncated Bergman ensemble have marked differences. While the Ginibre eigenvalues fill the disk as $n \to \infty$, the truncated Bergman points concentrate near $|z| = 1$ (*i.e.*, $\frac{1}{n} K_n(z, \bar z)$ tends weakly to $\delta_{|z|=1}$). Once again, the analytic CLT only "sees the boundary".

The proof is largely inspired by the ideas of Diaconis-Evans [6] where the analogous moment formula is established for the eigenvalues of the Haar distributed unitary group. Those points yield yet another example in the above class: there $d\mu_n(z) = \delta_{|z|=1}$ and $\rho = 1$. Also in that case, [6] shows the equality (30) to hold for finite n as soon as $n \ge (\sum_{j=1}^{k} j a_j) \vee (\sum_{j=1}^{k} j b_j)$. The basic observation is that the integrand on the left hand side of (30) is comprised of symmetric polynomials in the points $z_1, \dots, z_n$; one would like to to expand this object in a convenient basis.

Let A_n be the vector space of symmetric polynomials in the variables $z^{\oplus n} = z_1, ..., z_k$ of degree at most n. Let Λ_n denote the set of partitions of integers at most n. Given a partition $\lambda = (\lambda_1 \ge \lambda_2 \ge \cdots \ge ...) \in \Lambda_n$, bring in the corresponding *Schur function*,

$$s_\lambda(z^{\oplus n}) = \frac{\det\Big(z_k^{\lambda_\ell + n - \ell}\Big)_{k,\ell=1}^{n}}{\det\Big(z_k^{n-\ell}\Big)_{k,\ell=1}^{n}}, \tag{31}$$

and recall the following well known facts.

Theorem 16. *([16], Chapter 1) The Schur functions $s_\lambda(z^{\oplus n}), \lambda \in \Lambda_n$ form a basis for A_n. Consider the inner product $\langle \cdot, \cdot \rangle$ that makes them an orthonormal basis. For $f_n(z^{\oplus n}) = \prod_{j=1}^k p_k^{a_k}(z^{\oplus n})$ and $g_n(z^{\oplus n}) = \prod_{j=1}^k p_k^{b_k}(z^{\oplus n}) \in A_n$, which are products of simple power sum functions, it holds that*

$$\langle f_n, g_n \rangle = \delta_{ab}(\prod_{j=1}^{n} j^{a_j} a_j!). \tag{32}$$

The inner products are compatible as n varies in the sense that if $m > n$, and $\lambda \in \Lambda_n$ we have $\langle f_n, s_\lambda(z^{\oplus n}) \rangle = \langle f_m, s_\lambda(z^{\oplus m}) \rangle$.

Lemma 17. *For Schur functions in the points of (K, μ_n) as above,*

$$\mathbf{E}_n\Big[s_\lambda(z^{\oplus n})\,\overline{s_\pi(z^{\oplus n})}\Big] = \delta_{\lambda\pi} \prod_{\ell=0}^{n-1} \frac{M(n, 2(n+\lambda_\ell - \ell))}{M(n, 2(n-\ell))}. \tag{33}$$

Proof. The denominator in (31) is a Vandermonde determinant, and thus the interaction term (see (1)) in the integral of question is cancelled:

$$\mathbf{E}_n\Big[s_\lambda(z^{\oplus n})\,\overline{s_\pi(z^{\oplus n})}\Big] = \frac{1}{\mathcal{Z}_n} \int_{\mathbb{C}} \cdots \int_{\mathbb{C}} \det\Big(z_k^{\lambda_\ell + n - \ell}\Big) \det\Big(z_k^{\pi_\ell + n - \ell}\Big) d\mu_n(z_1) \ldots d\mu_n(z_n).$$

The normalizer here is $\mathcal{Z}_n = n! \prod_{\ell=1}^n M(n, 2\ell)$. Now, expanding each determinant on the right hand side, the generic term we get is a constant multiple of

$$\prod_{k=1}^{n} \int_{\mathbb{C}} z^{\lambda_k - k} \bar{z}^{\pi_{\sigma_k} - \sigma_k} |z|^{2n} \, d\mu_n(z)$$

with a permutation σ. Since λ and π are monotone, this will vanish by orthogonality of z and $\bar{z}$ when $\lambda \neq \pi$. When $\lambda = \pi$, there is a non-zero contribution from the diagonal, $\sigma = id$. To conclude, we compute

$$\mathbf{E}_n\Big[s_\lambda(z^{\oplus n})\,\overline{s_\lambda(z^{\oplus n})}\Big] = \frac{n!}{\mathcal{Z}_n} \prod_{k=0}^{n-1} \int_{\mathbb{C}} |z|^{2(\lambda_\ell + n - \ell)} d\mu_n(z),$$

but this is exactly (33). ☐

Proof of Theorem 14. Let $f_N = \prod_{j=1}^k p_j(z^{\oplus n})^{a_j}$ and $g_n = \prod_{j=1}^k p_j(z^{\oplus n})^{b_j}$, and let n_0 be at least of the degrees of f and g (this is independent of n). Now expand f_n, g_n with respect to the basis given by the Schur functions, and compute the expectation term by term. This expansion is finite, and the coefficients $\langle f_n, s_\lambda(z^{\oplus n}) \rangle$, $\langle g_n, s_\lambda(z^{\oplus n}) \rangle$ do not depend on n for $n \geq n_0$ by Theorem 16.

Since (33) has a bounded number of factors that are not 1, our assumption and Lemma 17 implies that for λ, ν fixed we have $\mathbf{E}_n[s_\lambda \bar{s_\nu}] \to \delta_{\lambda\nu}$. Thus we have $\mathbf{E}_n[f_n \bar{g}_n] \to \langle f_{n_0}, g_{n_0} \rangle$. The claim now follows by (32) and the moment formula for complex Gaussians. $\square$

Proof of Corollary 15. For analytic $f(z)$, $f(0) = \frac{1}{\pi}\int_{\mathbb{U}} f(z) d^2 z = \frac{1}{2\pi}\int_{\partial\mathbb{U}} f(z) dz$, which explains the centralizer. While for the Ginibre ensemble, we could simply quote the result of Section 7, the analyticity allows for a simpler approach amenable to more general ensembles subject perhaps to additional conditions in the spirit of (29). Preferring to be concrete (and brief), we restrict ourselves to the truncated Bergman case.

Rather than using the dbar representation, the variance in the analytic case may be computed by the more familiar Cauchy integral formula: with the reference measure now uniform on the unit disk,

$$\frac{1}{4\pi^2}\int_{C_\delta}\int_{C_\delta} f(\nu)\overline{f(\eta)}\Big[\int_{\mathbb{U}} \frac{K_n(z,\bar{z})}{(z-\nu)(\bar{z}-\bar{\eta})} d^2 z - \int_{\mathbb{U}}\int_{\mathbb{U}} \frac{|K_n(z,\bar{w})|^2}{(z-\nu)(\bar{z}-\bar{\eta})} d^2 z d^2 w\Big] d\nu d\bar{\eta}.$$

Here C_δ is a circle of radius $1+\delta > 1$ (about the origin) within the region of analyticity of f. What to do next is plain:

$$\int_{\mathbb{U}} \frac{K_n(z,\bar{z})}{(z-\nu)(\bar{z}-\bar{\eta})} d^2 z = \sum_{\ell=0}^{\infty} \frac{1}{(\nu\bar{\eta})^{\ell+1}} \sum_{m=0}^{n-1} \frac{m+1}{m+\ell+1},$$

and

$$\int_{\mathbb{U}}\int_{\mathbb{U}} \frac{|K_n(z,\bar{w})|^2}{(z-\nu)(\bar{w}-\bar{\eta})} d^2 z d^2 w = \sum_{\ell=0}^{n-1} \frac{1}{(\nu\bar{\eta})^{\ell+1}} \sum_{m=0}^{n-\ell-1} \frac{m+1}{m+\ell+1},$$

after expanding both the kernel and the functions $(z-\cdot)^{-1}$, $(w-\cdot)^{-1}$ and integrating termwise. With $|\nu\eta| > 1$ there is enough control to pass the limit inside the summations and conclude that the variance tends to

$$\frac{1}{4\pi^2}\int_{C_\delta}\int_{C_\delta} f(\nu)\overline{f(\eta)}\Big(\sum_{\ell=1}^{\infty} \frac{\ell}{(\nu\bar{\eta})^{\ell+1}}\Big) d\nu d\bar{\eta} \;\; = \;\; \sum_{\ell=1}^{\infty} \ell \Big| \frac{1}{2\pi}\int_{\partial U} \frac{f(\nu)}{\nu^{\ell+1}} d\nu \Big|^2,$$

as $n \to \infty$. Here we have used the analyticity to pull the integral back to $\partial\mathbb{U}$ after the limit was performed. This last expression, and much of the method getting there, is now recognized from the Ginibre boundary case (Lemma 13). In particular, that it equals the advertised $\frac{1}{\pi}\int_{\mathbb{U}} |f'(z)|^2 d^2 z$ has already been explained. $\square$

Acknowledgements. The work of B.R. was supported in part by NSF grant DMS-0505680, and that of B.V. by a Sloan Foundation fellowship, by the Canada Research Chair program, and by NSERC and Connaught research grants. Both authors thank the hospitality of MSRI (spring 2005) where this project was initiated.

References

[1] Anderson, G. and Zeitouni, O. (2006). A CLT for a band matrix model *Probab. Theory and Related Fields* **134**, 283-338.

[2] Bai, Z.D. (1997). Circular Law. *Ann. Probab.* **25**, 494-529.

[3] Bai, Z.D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Annals of Probability* **32**, 533-605.

[4] Bell, S.R. *The Cauchy transform, potential theory, and conformal mapping.* Studies in Advanced Mathematics. CRC Press, Boca Raton, Fl, 1992.

[5] Costin, O. and Lebowitz, J. (1995). Gaussian fluctuations in random matrices. *Phys. Review Letters* **75**, 69-72.

[6] Diaconis, P. and Evans, S.N. (2001). Linear functionals of eigenvalues of random matrices. *Trans. AMS* **353**, 2615-2633.

[7] Forrester, P.J. (1999). Fluctuation formula for complex random matrices. *J. Phys. A: Math and General* **32**, 159-163.

[8] Ginibre, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6**, 440-449.

[9] Guionnet, A (2002). Large deviations and upper bounds for non-commutative functionals of Gaussian large random matrices. *Ann. Inst. H. Poincaré Probab. Statist.* **38**, 341-384.

[10] Hammersley, J. M The zeros of a random polynomial. *Proc. of the Third Berkeley Symposium on Mathematical Statistics and Probibility, 1954-1955, vol. II*, 89-111. University of California Press, Berkeley and Los Angeles, 1956.

[11] Hough, J. B., Krihnapur, M, Peres, Y., and Virág, B. (2006). Determinantal Processes and Independence. *Probability Surveys* **3**, 206-229, *arXiv:math.PR/0503110.*

[12] Johansson, K. (1997). On random matrices from the classical compact groups. *Ann. Math.* **145**, 519-545.

[13] Johansson, K. (1998). On the fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. Journal*, **91**, 151-204.

[14] Kostlan, E. (1992). On the spectra of Gaussian matrices. *Linear Algebra Appl.*, **162/164**, 385-388.

[15] Macchi, O. (1975). The coincidence approach to stochastic point processes. *Adv. Appl. Probab.*, **7**, 83-122.

[16] Macdonald, I.G. *Symmetric functions and Hall Polynomials*, Claredon Press, Oxford, 1979.

[17] Peres, Y. and Virág, B. (2005). Zeros of the i.i.d. Gaussian power series: a conformally invariant determinantal process. *Acta. Math.* **194**. 1 - 35.

[18] Rider, B., Silverstein, J.W. (2005). Gaussian fluctuations for non-Hermitian random matrix ensembles. *To appear, Ann. Probab., arXiv:math.PR/0502400.*

[19] Sheffield, S. (2005). Gaussian free fields for mathematicians. *Preprint, arXiv:math.PR/0312099.*

[20] Spitzer, F. (1956). A combinatorial lemma and its application to probability theory. *Trans. AMS* **82**, no. 2, 323-339.

[21] Spohn, H. Interacting Brownian particles: A study of Dyson's model, in *Hydrodynamic behavior and interacting particle systems,* G. Papanicolau, ed. Springer, New York, 1987.

[22] Soshnikov, A (2000). Determinantal random point fields. *Russian Math. Surveys* **55**, 923-975.

[23] Soshnikov, A (2000). Central Limit Theorem for local linear statistics in classical compact groups and related combinatorial identities. *Ann. Probab.* **28**, 1353-1370.

[24] Soshnikov, A (2002). Gaussian limits for determinantal random point fields. *Ann. Probab.* **30**, 171-181.

[25] Wieand, K. (2002). Eigenvalue distributions of random unitary matrices. *Probability Theory and Related Fields* **123**, 202-224.

Bálint Virág, Departments of Mathematics and Statistics, University of Toronto, ON, M5S 2E4, Canada.
`balint@math.toronto.edu`, `www.math.toronto.edu/~balint`

B. Rider, Department of Mathematics, University of Colorado at Boulder, Boulder, CO 80309.
brider@euclid.colorado.edu, math.colorado.edu/~brider